Building and Linking Unit Record Databases: Lessons from Experience

Peter Ewell
National Center for Higher Education Management Systems (NCHEMS)

Hans L’Orange
State Higher Education Executive Officers (SHEEO)

December 16, 2008
Guiding Principles

- Reinforce the capacity of states to conduct analysis that support state-level policymaking.
- Develop approaches that work regardless of who “owns” the data.
- Keep it simple!
- Allow for incremental development.
- Avoid “flash point” topics.
- Develop approaches in an open and transparent environment.
Sound Database Linking and Design

- Begins with Determining the Needs and Interests of Current and Potential Stakeholders
- Involves Careful Consideration of the Kinds of Questions that Potential Users Need to Answer
- Is Best Conducted Under Legislative Authority that Authorizes Links but Does Not Specify Exactly How They Should Be Accomplished
- Requires Secure Unique Identifiers, and Aligned Data Element Definitions and Code Structures
- Requires Accessible, Comprehensive, and Up-to-Date Documentation
**Connected - The Additional Information Available for Policy Decisions**

- Postsecondary performance by type of students, region, school district, etc.
- Impact of more rigorous course-taking
- Aligning K-12 assessments and postsecondary expectations/success
- High school feedback

**K-12**

- Demographics
- Course-Taking
- CSAP
- Additional Assessments
- Dropout and Graduation Rates

**Postsecondary**

- Demographics
- Remediation
- Course-Taking
- Retention Rates
- Graduation Rates
- Certificates and Degrees Awarded (by Field)

**Employment (UI)**

- Demographics
- Employment
- Employment Type
- Earnings

- Employment in state by type of students/graduates, region, college, etc.
- Employment in field of training (certain fields only)
- Return on investment - earnings as a result of postsecondary training
- Meeting state and regional workforce and economic needs
Some Important Policy Questions

- How Many (and Which) High School Graduates Enroll in College and Are Prepared for College-Level Work?

- What High-School Course-Taking Patterns and Performance Predict Postsecondary Success?

- How Many (and Which) College Entrants Persist and Earn Credentials [at the Institution Entered, Any Institution in State, or Anywhere]?

- How Many Graduates of the State’s Colleges and Universities Remain in the State and are Employed?

- How Successful are Adult Basic Education and Workforce Training Programs in Developing Employment and Literacy Skills Among Returning Adults?
Key Data Elements: K-12 System

- Demographics
- Completion, Graduation (Include GED as well)
- GPA and Class Standing (and Size)
- Particular Areas of Course-Taking (Math, Science, Upper English, AP) and Grades Earned
- Assessments Taken and Results (HS Exit, College Entrance, AP, Other)
- If Applicable, ABE and Other Basic Skills Training
Key Data Elements: Postsecondary

- Demographics
- Persistence and Credentials Earned (by Level and Field)
- Credit Loads, Credit Completion, GPA
- Need for Remediation (Reading, Writing, Math) and Remediation Participation
- Financial Aid Participation and Percentage Need Covered
- Particular Areas of Course-Taking and Grades Earned
- If Applicable, ABE and Other Basic Skills Training
Key Data Elements: Employment

- Demographics
- Employment Status (Hours Worked)
- Employment Field
- Earnings
- [Federal or Military Employment]
- [State Residence from DMV Records]
- [Social Service Usage]
Some Particular Questions About Linking

- Will the Resulting Enhanced Data be Permanent?

- Who “Owns” the Resulting Enhanced Data?
  - Who is Responsible for its Integrity?
  - What Conditions Govern its Reporting and Use?

- What are the Essential Contents of a Formal Data Sharing Agreement?

- How are the Provisions of a Data Sharing Agreement Enforced?
Making the Match: Best Practice Suggests...

- Using a Common Unique Identifier that is Not the SSN

- Using Multiple Additional Data Elements (DOB, Name, Gender)

- A Single Common Identifier or Routine (as Opposed to Bilateral Agency Agreements)

- A Third-Party Matchmaker that Cannot Independently Crack the Unique Identifier
Structuring the Match: Experience Suggests…

- Only About a Dozen Data Elements in Each Database will Account for 90% of the Action
- Hardware/Software Changes in Parent Systems Can Affect Implementation and Maintenance
- Data Quality is the Biggest Challenge
- Constant User Training will be Required
Options for Organizing the Relationships

- Fully Relational Structure Housing All Data Elements
- “Common Core” of Frequently-Used Data Elements Housed in a Separate, Mutually-Accessible Place
- Access to Limited Subsets of Data Elements Identified by Each Agency Using a Common Protocol
- Different Kinds of Access and Pathways Established by Bilateral Data Access Agreements Between Individual Agencies
A Case Study – Learning from Others

Background

- Higher Education agency has responsibility to report on higher education
- Labor has a legislative requirement to report on employment and wages
- History of sharing information between Higher Education and Labor
- Demonstrate importance of higher education in workforce development
A Case Study – Learning from Others
State Impetus

- Workforce investment is more than just getting people jobs
- Involves identifying the linkage between education and work
- Create greater efficiency in the labor market and yield economic benefits by
  - Enhancing education systems
  - Getting persons into relevant jobs sooner
A Case Study – Learning from Others
How Does it Happen?

- State Employment and Training Commission has primary responsibility for delivering an annual state report.
- State took a broad view of “training” to include higher education and wants to demonstrate value of higher education.
- Department of Labor collects and compiles data including employment outcomes and participant demographics.
- Synchronized with WIA performance measures.
A Case Study – Learning from others

Issues that needed to be addressed

- Administrators and technicians had problems with
  - Confidentiality and data sharing
  - Data definitions
  - Database and system changes
  - Data collection processes
A Case Study – Learning from Others Finding Solutions

- Employment and Training Commission sent out a decree making all program managers responsible for supplying data
- Administrators worked to resolve policy and procedural issues for data sharing
- Data technicians quietly put their heads together to solve data and programming issues
- Support and promotion by State Department of Higher Education and Department of Education (K-12)
- Labor’s Office of Research took the lead in finding compatible data definitions and processing solutions
A Case Study – Learning from Others Creating and Sharing Knowledge

- Each higher education system provides the Labor Department with a file containing student identifiers, program completion, and demographics.

- DOL does data matches and generates required reports.

- DHE’s accountability report and separately published *Higher Education and the Workforce* provide information on which graduates were employed after graduation by industry sector, race/ethnicity, gender, average wages, and company size.
A Case Study – Learning from Others
Structure of the Report

- Program summary (narrative)
- Services: Types and quantities delivered
- Demographics: Who is served
- Results: Employment outcomes
  - Participants and university completers (numbers)
  - Entered employment
  - Retained employment
  - Weekly wages and change in wages
A Case Study – Learning from Others
Policy Questions to be Explored

■ Are the state’s colleges producing enough graduates to meet current and future workforce needs?

■ Are there leaks in the college pipeline from high school to college to work?

■ What is the state’s return on investment in higher education in terms of the earnings of college graduates?